Analysis of Bidirectional Associative Memory using SCSNA and Statistical Neurodynamics

Hayaru Shouno and Shoji Kido
Dept. of Computer Science and Systems Engineering,
*Faculty of Engineering, Yamaguchi University**

Masato Okada
Brain Research Institute, RIKEN
(Dated: February 2, 2008)

Bidirectional associative memory (BAM) is a kind of an artificial neural network used to memorize and retrieve heterogeneous pattern pairs. Many efforts have been made to improve BAM from the the viewpoint of computer application, and few theoretical studies have been done. We investigated the theoretical characteristics of BAM using a framework of statistical-mechanical analysis. To investigate the equilibrium state of BAM, we applied self-consistent signal to noise analysis (SCSNA) and obtained a macroscopic parameter equations and relative capacity. Moreover, to investigate not only the equilibrium state but also the retrieval process of reaching the equilibrium state, we applied statistical neurodynamics to the update rule of BAM and obtained evolution equations for the macroscopic parameters. These evolution equations are consistent with the results of SCSNA in the equilibrium state.

I. INTRODUCTION

Bi-directional associative memory (BAM) [1] is a kind of an associative memory model which is an artificial neural network. The principle function of associative memory is to memorize multiple patterns and to retrieve the correct one when a pattern key is given.

Autocorrelation associative memory (AAM), sometimes called the Hopfield model [2], is also a kind of associative memory. AAM tries to retrieve a stored pattern when a degraded pattern is given as an association key; this type of retrieving is called homogeneous association. In contrast, BAM stores multiple pattern pairs and tries to retrieve a complete stored pattern pair when a degraded piece of the pair is given as an association key. Thus, BAM is called a heterogeneous pattern association model.

In the field of neural networks, many efforts have been made to improve BAM from the viewpoint of computer application [3] [4] [5] [6] [7] [8] [9] [10], and few theoretical analyses have been reported [11] [12] [13]. The theoretical analysis of BAM has evolved with a focus on storage capacity, which means how many patterns can be stored in a network consisted of N neural units. Yanai *et al.* suggested that BAM can be regarded as a variation of AAM in which connections are systematically reduced [12]. They also showed that the relative storage capacity, in which a finite amount of retrieval error is allowed, of BAM to be around $0.22N$. Haines & Hecht-Nielsen analyzed BAM along the same way, and reported its absolute capacity, in which no retrieval error is allowed, to be $O(N/\log N)$ [11]. Tanaka *et al.* analyzed BAM using a replica method (see [14]), which is a statistical-mechanical analysis method, and showed its relative capacity to be $0.1998N$ [13]. These analyses mainly focused on the equilibrium state of BAM, and the transient process of retrieving, which means how to reach the equilibrium state, was not so conducted. However, analysis of the retrieval process is as important as that of the equilibrium state.

In this paper, we have analyzed the equilibrium state of BAM using the self-consistent signal-to-noise analysis (SCSNA)[15]. We found that the relative capacity was $0.1998N$, which agrees with the result of Tanaka *et al.* We also investigated the retrieval process of BAM; we derived macroscopic dynamical equations using the statistical neurodynamics, which was theoretically derived in the same manner as the SCSNA [16] [17], and compared the results of between the statistical neurodynamics with those of computer simulation. Applying the statistical neurodynamics to BAM, we obtained the evolution equations for the macroscopic parameters. In the limit of these evolution equations, that is, macroscopic parameters of BAM reached

*Electronic address: shouno@ai.csse.yamaguchi-u.ac.jp

the equilibrium state, we found these values were consistent with the results of SCSNA. We also compared the results of applying the statistical neurodynamics with those of the computer simulation and obtained quantitative support for our analysis.

We describe the BAM formulation in Section II, and we apply the SCSNA and show the results of equilibrium state analysis in Section III. In Section IV, we derived the evolution equations of macroscopic parameters using the statistical neurodynamics and compared the results with those of computer simulation in Section V.

II. FORMULATION

As shown in Fig. 1, BAM is a two-layered neural network model [1]. The first layer consists of cN neural units ($c \sim O(1)$), and the state of the layer is denoted as $\boldsymbol{x}$ with the components denoted as x_i ($1 \leq i \leq cN$). The state of the second layer, which has $\tilde{c}N$ units, is denoted as $\tilde{\boldsymbol{x}}$, and the jth unit state is described as $\tilde{x}_j$ ($1 \leq j \leq \tilde{c}N$). Each layer is connected by interlayer connection $\boldsymbol{J}$ with the components described as J_{ij} ($1 \leq i \leq cN$, $1 \leq j \leq \tilde{c}N$). J_{ij} represents the connection weight between the first layer unit x_i and the second layer unit $\tilde{x}_j$.

We prepare p binary pattern pairs denoted as $\{\boldsymbol{\xi}^\mu, \tilde{\boldsymbol{\xi}}^\mu\}$ ($\mu = 1, \cdots, p$), where the superscript μ denotes the pattern pair index. Pattern vector $\boldsymbol{\xi}^\mu$ corresponds to the first layer, and $\tilde{\boldsymbol{\xi}}^\mu$ corresponds to the second layer. Thus $\boldsymbol{\xi}^\mu$ and $\tilde{\boldsymbol{\xi}}^\mu$ have cN and $\tilde{c}N$ components, respectively, and each component, which is described as ξ_i^μ and $\tilde{\xi}_j^\mu$ ($1 \leq i \leq cN$, $1 \leq j \leq \tilde{c}N$), respectively, is generated from uniform i.i.d.:

$$\mathrm{Prob}[\xi_i^\mu = \pm 1] = \frac{1}{2}, \tag{1}$$

$$\mathrm{Prob}[\tilde{\xi}_j^\mu \pm 1] = \frac{1}{2}. \tag{2}$$

Assuming the number of stored pattern pairs to be $p \sim O(N)$, we define a quantity $\alpha \left(= \frac{p}{N}\right)$, and use it for the loading rate.

To determine the interlayer connection weight J_{ij}, which connects x_i and $\tilde{x}_j$, we use a correlation-based learning rule:

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{\alpha N} \xi_i^\mu \tilde{\xi}_j^\mu. \tag{3}$$

All the pattern pair correlations between $\boldsymbol{\xi}^\mu$ and $\tilde{\boldsymbol{\xi}}^\mu$ are embedded in connection weight $\boldsymbol{J}$. In this notation, the connections are not symmetrical, that is $J_{ij} \neq J_{ji}$.

In the retrieving, we use a synchronous update rule for each layer; that is, all units in each layer are updated synchronously, and these layers are updated alternately. The rules for updating the ith unit in the first layer and the jth unit in the second layer are

$$x_i^{2t} = F(\sum_{j=1}^{\tilde{c}N} J_{ij} \tilde{x}_j^{2t-1}), \tag{4}$$

$$\tilde{x}_j^{2t+1} = F(\sum_{i=1}^{cN} J_{ij} x_i^{2t}), \tag{5}$$

where t means the one step time and $F(\cdot)$ means the output function. In these formulations, the retrieval process is carried out as follows. In a initial state, $t = 0$, Association key $\boldsymbol{x}^0 (= \{x_i^0\})$ is given to the first layer. Then, all of the second layer units, $\tilde{x}_j^1$ ($1 \leq j \leq \tilde{c}N$), are updated using eq. (5), and the state of the second layer is described as $\tilde{\boldsymbol{x}}^1$. Next, $t = 1$, all the units in the first layer, x_i^2 ($1 \leq i \leq cN$), are updated using eq.(4), and the state is described as $\boldsymbol{x}^2$. After that the second layer is updated by eq.(5). For each $t = 2, 3, \cdots$, the alternate updating of each layer are carried out in the same way, and each layer state is denoted as $\boldsymbol{x}^{2t}$ and $\tilde{\boldsymbol{x}}^{2t+1}$.

To apply S/N analysis, we introduce overlaps, which means similarities between patterns. The overlaps between first layer state $\boldsymbol{x}^{2t}$ and the μth pattern, $\boldsymbol{\xi}^{\mu}$, and between second-layer state $\tilde{\boldsymbol{x}}^{2t-1}$ and $\tilde{\boldsymbol{\xi}}^{\mu}$ are described as follows, respectively:

$$m_{\mu}^{2t} = \frac{1}{cN}\sum_{i=1}^{cN} x_i^{2t}\xi_i^{\mu}, \tag{6}$$

$$\tilde{m}_{\mu}^{2t+1} = \frac{1}{\tilde{c}N}\sum_{j=1}^{\tilde{c}N} \tilde{x}_j^{2t+1}\tilde{\xi}_j^{\mu}. \tag{7}$$

Following the S/N analysis, we decomposed the inner term of $F(\cdot)$ in eqs.(4) and (5) into signal and noise components. Assuming the first pattern pair, $\{\boldsymbol{\xi}^1, \tilde{\boldsymbol{\xi}}^1\}$, is retrieved, the terms including overlaps m_1 and $\tilde{m}_1$, are signal components, *i.e.* m_1, $\tilde{m}_1 \sim O(1)$. Using these overlaps, eqs.(4) and (5) can be described as

$$x_i^{2t} = F(\tilde{c}\tilde{m}_1^{2t-1}\xi_i^1 + z_i^{2t-1}), \tag{8}$$

$$\tilde{x}_j^{2t+1} = F(cm_1^{2t}\tilde{\xi}_j^1 + \tilde{z}_j^{2t}), \tag{9}$$

where z_i^{2t-1}, and $\tilde{z}_j^{2t}$ are called as crosstalk noises, which prevents the target pair $\{\boldsymbol{\xi}^1, \tilde{\boldsymbol{\xi}}^1\}$ to be retrieved. These crosstalk noises are denoted

$$z_i^{2t-1} = \frac{1}{N}\sum_{\mu=2}^{\alpha N}\sum_{j=1}^{\tilde{c}N} \xi_i^{\mu}\tilde{\xi}_j^{\mu}\tilde{x}_j^{2t-1}, \tag{10}$$

$$\tilde{z}_j^{2t} = \frac{1}{N}\sum_{\mu=2}^{\alpha N}\sum_{i=1}^{cN} \tilde{\xi}_j^{\mu}\xi_i^{\mu}x_i^{2t}. \tag{11}$$

III. EQUILIBRIUM STATE ANALYSIS BY SCSNA

To derive equilibrium state macroscopic parameters, we use SCSNA[15], which is an extension of a naive signal-to-noise (S/N) analysis. Since the SCSNA treats the equilibrium states of an associative memory model, we omit the index t in the update rules. Hence we can rewrite eqs. (8) and (9) as

$$x_i = F(\tilde{c}\tilde{m}_1\xi_i^1 + z_i), \tag{12}$$

$$\tilde{x}_j = F(cm_1\tilde{\xi}_j^1 + \tilde{z}_j), \tag{13}$$

respectively.

In the SCSNA, the crosstalk noise term is decomposed into a systematic bias term and a Gaussian noise term with 0 mean[15]. The detailed formulas of SCSNA are described in appendix. We derive self-consistent

equations called order parameter equations. The following are the order parameter equations of BAM.

$$Y = F(\tilde{c}\tilde{m}\xi + \frac{\alpha\tilde{c}\tilde{U}}{1-c\tilde{c}U\tilde{U}}Y + \sqrt{\alpha r}z), \tag{14}$$

$$\tilde{Y} = F(cm\tilde{\xi} + \frac{\alpha cU}{1-c\tilde{c}U\tilde{U}}\tilde{Y} + \sqrt{\alpha\tilde{r}}z), \tag{15}$$

$$m = \int Dz\, \langle \xi Y \rangle_{\xi}, \tag{16}$$

$$\tilde{m} = \int Dz \langle \tilde{\xi}\tilde{Y} \rangle_{\tilde{\xi}}, \tag{17}$$

$$q = \int Dz \langle Y^2 \rangle_{\xi}, \tag{18}$$

$$\tilde{q} = \int Dz \langle \tilde{Y}^2 \rangle_{\tilde{\xi}}, \tag{19}$$

$$U = \frac{1}{\sqrt{\alpha r}} \int Dzz \langle Y \rangle_{\xi}, \tag{20}$$

$$\tilde{U} = \frac{1}{\sqrt{\alpha \tilde{r}}} \int Dzz \langle \tilde{Y} \rangle_{\tilde{\xi}}, \tag{21}$$

$$r = \frac{\tilde{c}}{(1-c\tilde{c}U\tilde{U})^2}(\tilde{q} + c\tilde{c}\tilde{U}^2 q), \tag{22}$$

$$\tilde{r} = \frac{c}{(1-c\tilde{c}U\tilde{U})^2}(q + c\tilde{c}U^2 \tilde{q}). \tag{23}$$

These equations are described in the manner of Shiino and Fukai [15]. Y and $\tilde{Y}$ represents the effective outputs for x_i and $\tilde{x}_j$, respectively. The stochastic variables ξ and $\tilde{\xi}$, obeying eq.(1) and (2), corresponds to a retrieving pattern components ξ_i^1 and $\tilde{\xi}_j^1$, and order parameters m and $\tilde{m}$ corresponds to overlaps m_1 and $\tilde{m}_1$. Note that the operators $\langle\cdot\rangle_{\xi}$ and $\langle\cdot\rangle_{\tilde{\xi}}$ mean the expectations for stochastic variables ξ or $\tilde{\xi}$, respectively. Each arguments of the function $F(\cdot)$ consists of three parts. The first terms, $\tilde{c}\tilde{m}\xi$ and $cm\tilde{\xi}$, come from the signal components, the second terms, $\frac{\alpha\tilde{c}\tilde{U}}{1-c\tilde{c}U\tilde{U}}Y$ and $\frac{\alpha cU}{1-c\tilde{c}U\tilde{U}}\tilde{Y}$, mean the systematic bias of the crosstalk noises (z_i, $\tilde{z}_j$) in eqs. (10) and (11), and each third term is assigned to be a Gaussian distribution with 0 mean and αr or $\alpha\tilde{r}$ variance. We solved the order parameter equations from (14) to (23) numerically and compared the results with those of simulations. Fig. 2. shows the equilibrium overlap m against the capacity parameter α. An overlap of 1 means that the BAM retrieves a stored pattern pair successfully. We obtained a relative capacity, α_c of 0.1998 in which the nontrivial solution $m \neq 0$ and $\tilde{m} \neq 0$ is disappeared. This agrees with the results of Tanaka *et al.* ($\alpha_c = 0.1998$), obtained with the replica method [13]. In fig.2, we show the simulation results as error-bars, which mean medians and quartile deviations for ten trials. The SCSNA results quantitatively explained the simulation results very well.

IV. RETRIEVAL PROCESS OF BAM

As we have seen, the SCSNA described the equilibrium state of BAM quantitatively. In this section, we consider a retrieval process of BAM, which means the transient process reaching the equilibrium state. The statistical neurodynamics, which is a theory for the retrieval for associative memory model, is based on S/N analysis. Amari and Maginu proposed a statistical neurodynamical theory on the S/N analysis[16], It was known that the storage capacity obtained by Amari & Maginu theory does not coincide with the results of the replica theory[18], and the size of the basin of attraction derived from Amari & Maginu theory is larger than the results of the computer simulation. Okada extended the Amari & Maginu theory to improve to resolve these difficulties[17], and obtained a macroscopic equation which has hierarchical structure. In the macroscopic equation, the first-order approximation corresponds to the Amari & Maginu theory, and the higher order approximation coincide with the replica theory.

For applying the statistical neurodynamics to BAM, we evaluate the crosstalk noises (eqs.(10) and (11)) in eqs. (8) and (9). Assuming the first pattern pair, $\{\boldsymbol{\xi}^1, \tilde{\boldsymbol{\xi}}^1\}$, is retrieved, we can regard the overlaps of

other pattern pairs, $\{m_\mu^{2t}, \tilde{m}_\mu^{2t+1}\}$ where $\mu \geq 2$, as small. Thus, we expand the state x_i^{2t} and $\tilde{x}_j^{2t+1}$:

$$x_i^{2t} = x_i^{2t,(\mu)} + \tilde{c}\tilde{m}_\mu^{2t-1}\xi_i^\mu F'(\tilde{c}\sum_{\nu\neq\mu}^{\alpha N}\tilde{m}_\mu^{2t-1}\xi_i^\nu), \tag{24}$$

$$\tilde{x}_j^{2t+1} = \tilde{x}_j^{2t+1,(\mu)} + cm_\mu^{2t}\tilde{\xi}_j^\mu F'(c\sum_{\nu\neq\mu}^{\alpha N} m_\nu^{2t}\tilde{\xi}_j^\nu), \tag{25}$$

for $\mu \geq 2$, where $\tilde{x}_j^{2t+1,(\mu)} = F(c\sum_{\nu\neq\mu}^{\alpha N} m_\nu^{2t}\tilde{\xi}_j^\nu)$ and $x_i^{2t,(\mu)} = F(\tilde{c}\sum_{\nu\neq\mu}^{\alpha N}\tilde{m}_\mu^{2t-1}\xi_i^\nu)$. Substituting eqs. (24) and (25) into eqs.(10) and (11), we obtain,

$$z_i^{2t+1} = \alpha\tilde{c}\tilde{U}_{2t+1}x_i^{2t,(\mu)} + Z_i^{2t+1}, \tag{26}$$

$$\tilde{z}_j^{2t} = \alpha c U_{2t}\tilde{x}_j^{2t-1,(\mu)} + \tilde{Z}_j^{2t}, \tag{27}$$

where

$$Z_i^{2t+1} = \frac{1}{N}\sum_{\mu=2}^{\alpha N}\sum_{j=1}^{\tilde{c}N}\xi_i^\mu\tilde{\xi}_j^\mu\tilde{x}_j^{2t+1,(\mu)} + \frac{\tilde{c}\tilde{U}_{2t+1}}{N}\sum_{\mu=2}^{\alpha N}\sum_{k\neq i}^{cN}\xi_i^\mu\xi_k^\mu x_k^{2t,(\mu)} + \tilde{c}c\tilde{U}_{2t+1}U_{2t}Z_i^{2t-1}, \tag{28}$$

$$\tilde{Z}_j^{2t} = \frac{1}{N}\sum_{\mu=2}^{\alpha N}\sum_{i=1}^{cN}\tilde{\xi}_j^\mu\xi_i^\mu x_i^{2t,(\mu)} + \frac{cU_{2t}}{N}\sum_{\mu=2}^{\alpha N}\sum_{l\neq j}^{\tilde{c}N}\tilde{\xi}_j^\mu\tilde{\xi}_l^\mu\tilde{x}_l^{2t-1,(\mu)} + c\tilde{c}U_{2t}\tilde{U}_{2t-1}\tilde{Z}_j^{2t-2}, \tag{29}$$

where

$$\tilde{U}_{2t+1} = \frac{1}{\tilde{c}N}\sum_{j=1}^{\tilde{c}N}F'(c\sum_{\nu\neq\mu}^{\alpha N}m_\nu^{2t}\tilde{\xi}_j^\nu), \tag{30}$$

$$U_{2t} = \frac{1}{cN}\sum_{i=1}^{cN}F'(\tilde{c}\sum_{\nu\neq\mu}^{\alpha N}\tilde{m}_\mu^{2t-1}\xi_i^\nu). \tag{31}$$

Since $x_i^{2t,(\mu)}$ and $\tilde{x}_j^{2t-1,(\mu)}$ are almost independent with ξ_i^μ and $\tilde{\xi}_j^\mu$, respectively, each Z_i^{2t+1} and $\tilde{Z}_j^{2t}$ can be regarded as independent identical Gaussian distributions, that is $Z_i^{2t+1} \sim N(0, \alpha r_{2t+1})$ and $\tilde{Z}_j^{2t} \sim N(0, \alpha\tilde{r}_{2t})$. Each noise variance, $E[(Z_i^{2t+1})^2] = \alpha r_{2t+1}$ and $E[(\tilde{Z}_j^{2t})^2] = \alpha\tilde{r}_{2t}$, can be described as

$$\begin{aligned}\alpha r_{2t+1} = {} & \alpha\tilde{c}\tilde{q}_{2t+1} + \alpha c\tilde{c}^2\tilde{U}_{2t+1}^2 q_{2t} + \alpha(\tilde{c}c\tilde{U}_{2t+1}U_{2t})^2 r_{2t-1} \\ & + 2c\tilde{c}\tilde{U}_{2t+1}U_{2t}E\left[Z_i^{2t-1}\frac{1}{N}\sum_{\mu=2}^{\alpha N}\sum_{j=1}^{\tilde{c}N}\xi_i^\mu\tilde{\xi}_j^\mu\tilde{x}_j^{2t+1,(\mu)}\right],\end{aligned} \tag{32}$$

$$\begin{aligned}\alpha\tilde{r}_{2t} = {} & \alpha c q_{2t} + \alpha\tilde{c}c^2 U_{2t}^2\tilde{q}_{2t-1} + \alpha(c\tilde{c}U_{2t}\tilde{U}_{2t-1})^2\tilde{r}_{2t-2} \\ & + 2c\tilde{c}U_{2t}\tilde{U}_{2t-1}E\left[\tilde{Z}_j^{2t-2}\frac{1}{N}\sum_{\mu=2}^{\alpha N}\sum_{i=1}^{cN}\tilde{\xi}_j^\mu\xi_i^\mu x_i^{2t,(\mu)}\right],\end{aligned} \tag{33}$$

where

$$\tilde{q}_{2t+1} = \frac{1}{\tilde{c}N}\sum_{j=1}^{\tilde{c}N}\left(\tilde{x}_j^{2t+1,(\mu)}\right)^2, \tag{34}$$

$$q_{2t} = \frac{1}{cN}\sum_{i=1}^{cN}\left(x_i^{2t,(\mu)}\right)^2. \tag{35}$$

The last terms in eqs.(32) and (33) are determined by correlations between the current state $\tilde{x}_j^{2t+1}$ and the previous state noise variable Z_i^{2t-1}, and between x_i^{2t} and $\tilde{Z}_j^{2t-2}$, respectively. Assuming that the $(n+1)$ previous state noise variables $Z_i^{2(t-n)-1}$ and $\tilde{Z}_j^{2(t-n)-2}$ have no correlation with the current state $\tilde{x}_j^{2t+1}$ and x_i^{2t}, respectively, we can expand r_{2t+1} and $\tilde{r}_{2t}$ as recurrence formulas:

$$\begin{aligned} r_{2t+1} &= \tilde{c}\tilde{q}_{2t+1} + c(\tilde{c}\tilde{U}_{2t+1})^2 q_{2t} + (c\tilde{c}\tilde{U}_{2t+1}U_{2t})^2 r_{2t-1} \\ &+ 2\tilde{c}\sum_{\eta=1}^{n}(c\tilde{c})^{\eta}\tilde{q}_{2t+1,2(t-\eta)+1}\prod_{\tau=t-\eta+1}^{t}\tilde{U}_{2\tau+1}U_{2\tau} \\ &+ 2c(\tilde{c}\tilde{U}_{2t+1})^2\sum_{\eta=1}^{n-1}(c\tilde{c})^{\eta}q_{2t,2(t-\eta)}\prod_{\tau=t-\eta+1}^{t}U_{2\tau}\tilde{U}_{2\tau-1}, \end{aligned} \tag{36}$$

$$\begin{aligned} \tilde{r}_{2t} &= cq_{2t} + \tilde{c}(cU_{2t}^2)^2\tilde{q}_{2t-1} + (\tilde{c}cU_{2t}\tilde{U}_{2t-1})^2\tilde{r}_{2t-2}, \\ &+ 2c\sum_{\eta=1}^{n}(\tilde{c}c)^{\eta}q_{2t,2(t-\eta)}\prod_{\tau=t-\eta+1}^{t}U_{2\tau}\tilde{U}_{2\tau-1} \\ &+ 2\tilde{c}(cU_{2t})^2\sum_{\eta=1}^{n-1}(\tilde{c}c)^{\eta}\tilde{q}_{2t-1,2(t-\eta)-1}\prod_{\tau=t-\eta+1}^{t}\tilde{U}_{2\tau-1}U_{2\tau-2}, \end{aligned} \tag{37}$$

where $\tilde{q}_{2t+1,2(t-n)+1}$ means a cross-correlation between the current state $\tilde{x}_j^{2t+1}$ and the n-step previous state $\tilde{x}_j^{2(t-n)+1}$, and $q_{2t,2(t-n)}$ means a cross-correlation between x_i^{2t} and $x_i^{2(t-n)}$. These variables can be also described with the macroscopic parameters across the n-step previous state. The complete formula is described in the appendix.

We obtain the evolution equations for macroscopic parameters as follows:

$$Y^{2t} = F(\tilde{c}\tilde{m}_{2t-1}\xi + \sqrt{\alpha r_{2t-1}}z), \tag{38}$$

$$\tilde{Y}^{2t+1} = F(cm_{2t}\tilde{\xi} + \sqrt{\alpha\tilde{r}_{2t}}z), \tag{39}$$

$$m^{2t} = \int Dz\langle\xi Y^{2t}\rangle_{\xi}, \tag{40}$$

$$\tilde{m}^{2t+1} = \int Dz\langle\tilde{\xi}\tilde{Y}^{2t+1}\rangle_{\tilde{\xi}}, \tag{41}$$

$$q_{2t} = \int Dz\langle(Y^{2t})^2\rangle_{\xi}, \tag{42}$$

$$\tilde{q}_{2t+1} = \int Dz\langle(\tilde{Y}^{2t+1})^2\rangle_{\tilde{\xi}}, \tag{43}$$

$$U_{2t} = \frac{1}{\sqrt{\alpha r_{2t-1}}}\int Dzz\langle Y^{2t}\rangle_{\xi}, \tag{44}$$

$$\tilde{U}_{2t+1} = \frac{1}{\sqrt{\alpha\tilde{r}_{2t}}}\int Dzz\langle\tilde{Y}^{2t+1}\rangle_{\tilde{\xi}} \tag{45}$$

In these order parameter equations, Y^{2t} and $\tilde{Y}^{2t+1}$ correspond to the x_i^{2t} and $\tilde{x}_j^{2t+1}$, respectively. The overlaps for the first pattern pair, m_1^{2t} and $\tilde{m}_1^{2t+1}$, which mean retrieval degree, correspond to the m^{2t} and $\tilde{m}^{2t+1}$, respectively.

Yanai *et al.* [12] applied the one-step analysis, which corresponds to Amari & Maginu theory. In their analysis, the macroscopic order parameter equations for Y_i^{2t}, $\tilde{Y}_j^{2t+1}$, m_1^{2t}, $\tilde{m}_1^{2t+1}$, q_{2t}, $\tilde{q}_{2t+1}$, U_{2t},$\tilde{U}_{2t+1}$, which are described in eqs. from (38) to (45), are identical to those of our analysis. The differences are in evaluating of noise variances, that is, r_{2t+1} and $\tilde{r}_{2t}$. They ignored the noise correlation, and derived these values as

$$r_{2t+1} = \tilde{c}\tilde{q}_{2t+1} + c(\tilde{c}\tilde{U}_{2t+1})^2 q_{2t} \tag{46}$$

$$\tilde{r}_{2t} = cq_{2t} + \tilde{c}(cU_{2t}^2)^2\tilde{q}_{2t-1}. \tag{47}$$

In their result, the critical capacity α_c is equal to 0.27, which is not equal to our SCSNA analysis and the replica analysis ($\alpha_c = 0.1998$ for both analyses). This overestimation comes from the lack of noise correlation evaluation.

In our analysis, we consider the effect of crosstalk noise correlation across n-step previous state, and obtained the eqs.(32) and (33) which includes Yanais' analysis (eqs.(46) and (47)). In the next section, we show that the analysis accuracy improves as n in increased ($n = 2, 3, \cdots$). Hereafter, we call the statistical neurodynamics considering across the n-step previous state as the "n-step" analysis in the following. "Full-step" analysis means using all the macroscopic parameters from the initial state ($t = 0$) to the current state.

V. RESULT

In this section, we compare the results of the statistical neurodynamics with those of computer simulation. Fig. 3 shows the time evolution of the overlap m_1^{2t}, which means how well the pattern $\boldsymbol{\xi}^1$ is retrieved in the first layer at $2t$. Each abscissa axis represents the time step t and each ordinate axis describes overlap m_1^{2t}. Convergence of the overlap m_1^{2t} to 1.0 means success in retrieving $\boldsymbol{\xi}^1$. In the graphs, we show several evolution curves in which the initial overlap (m_1^0) starts with a different state.

Fig. 3(a) shows the simulation results. We set the number of neurons N to $10,000$, $c = \tilde{c} = 1$, and the number of stored pattern pairs was indexed as $\alpha = 0.15$. The retrieval was successful when we set the initial overlap larger than 0.4, and it failed when we set it to 0.3 or less. Fig. 3(b) shows the results of the one-step analysis, and figs.3(c) to (e) shows the result of the 2-step, 3-step, and full step analysis, respectively. In each analysis result, the overlap converged to 0 when retrieving failed because of assuming infinite neuron units ($N \to \infty$) exist. In the simulation results, fig.3(a), the system settled into a spurious memory state when retrieving failed because the number of units was finite ($N = 10,000$). Therefore, the curves starting at 0.1 to 0.3 can be regarded as retrieval failures.

As shown in fig. 3(b), the one-step analysis says that retrieving is successful when the initial overlap is 0.3, which does not agree with the simulation results. Fig. 3(c) shows the results for $n = 2$, *i.e.* the 2-step analysis. The 2-step analysis says that the retrieval is a failure when the initial overlap is 0.3, which agrees with the simulation results. Fig. 3(d) and (e) show the 3-step results and the full-step analysis results, respectively. Each figure shows similar characteristics, and the results agree with those of the simulation results shown as fig. 3(a). Since the 3-step analysis results are very similar to the full-step analysis results the 3-step analysis is enough for approximating the full-step analysis. In other words, the previous 3-step correlations are effective for BAM retrieval.

In the statistical neurodynamics analysis, the equilibrium state is described as the limit of the transient process, and the order parameters should be consistent to the result of the SCSNA. Fig. 4 shows the memory capacity and the basin of attraction , which means the degrading limit of the retrievable pattern in the initial state measured by the overlap m_1^0. For example, in fig. 3(d), the retrieval is successful when starting at $m_1^0 = 0.4$, while it is a failure when starting at $m_1^0 = 0.3$. There is thus a basin when m_1^0 is between 0.3 and 0.4 for $\alpha = 0.15$. The dashed curves in fig.4 are derived from the statistical neurodynamicses. In these curves, the upper part shows equilibrium overlap m_1^∞ in successful retrieval and the lower part shows the basin of attraction m_1^0. When we set the initial overlap m_1^0 to be in the area surrounded by these curves, the retrieval will be success. Therefore, the area surrounded by these curves represents the successful retrieval area. It is clear that the one-step analysis overestimates both the relative capacity and the basin of attraction[12]. The theoretical estimation accuracy improves and comes close to that of the SCSNA analysis asymptotically as the analysis accuracy is improved (2-step, 3-step, $\cdots$.) We also show the basin derived from the simulation results using error-bars in fig.4. The results of simulation agree with those of the statistical neurodynamics quantitatively.

VI. CONCLUSION

We derived the macroscopic parameters of BAM in the equilibrium state by using the SCSNA and obtained the critical capacity α_c as 0.1998. The results agreed with the previous results and the simulation results.

We also analyzed the transient process of BAM using the statistical neurodynamics and obtained the evolution equations for the macroscopic parameters. Comparison of the numerical solutions with the simulation

results, we showed that the analysis results can explain the simulation results with sufficient accuracy for the transient process. Therefore, to explain the transient process of BAM quantitatively, it is sufficient to consider the 3-step statistical neurodynamics, which means that the crosstalk noise has effective correlation across the 3-step previous state.

[1] B. Kosko. Bidirectional Associative Memory. *IEE Trans. Systems, Man and Cybernetics*, 18:49–60, 1988.

[2] J. J. Hopfield and D. W. Tank. Computing with Neural Circuits: A Model. *Science*, 233:625–633, 1986.

[3] M. H. Hassoun and A. M. Youssef. A High-Performance Recording Algorithm for Hopfield Model Associative Memories. *Optical Engineering*, 27(1):46–54, 1989.

[4] P. K. Simpson. Higher-Ordered and Intraconnected Bidirectional Associative Memories. *IEEE Trans. on Systems, Man, Cybernetics*, 20(3):637–652, 1990.

[5] Y. F. Wang, J. B. Cruz, and J. H. Mulligan. Two Coding Strategies for Bidirectional Associative Memory. *IEEE Trans. on Neural Networks*, 1(1):81–92, 1991.

[6] X. Zhuang, Y. Huang, and S. S. Chen. Better Learning for Bidirectional Associative Memory. *IEEE Trans. on Neural Networks*, 6(8):1131–1146, 1993.

[7] H. Oh and S. C. Kothari. Adaptation of the Relaxation Method for Learining in Bidirectional Associative Memory. *IEEE Trans. on Neural Networks*, 5(4):573–583, 1994.

[8] C. C. Wang and H. S. Don. An analysis of high-capacity discrete exponential bam. *IEEE Trans. on Neural Networks*, 6(2):492–496, 1995.

[9] Z. Wang. A Bidirectional Associative Memory Based on Optimal Linear Associative Memory. *IEEE Trans. on Computers*, 45(10):1171–1179, 1996.

[10] H. Shi, Y. Zhao, and X. Zhuang. A General Model for Bidirectional Associative Memories. *IEEE Trans. on Systems, Man and Cybernetics-Part B: Cybernetics*, 28(4):511–519, 1998.

[11] K. Haines and R. Hecht-Nielsen. A BAM with Increased Information Storage Capacity. In *Proc. IEEE Conf. Neural Networks*, volume 1, pages I–181–190, San Diego, 1988.

[12] H. F. Yanai, Y. Sawada, and S. Yoshizawa. Dynamics of an Auto-Associative Neural Network Model with Arbitrary Connectivity and Noise in the Threshold. *Network*, 2:295–314, 1991.

[13] T. Tanaka, S. Kakiya, and Y. Kabashima. Capacity Analysis of Bidirectional Associative Memory. In *Proc. Seventh Int. Conf. Neural Information Processing*, volume 2, pages 779–784, Taejon, Korea, 2000.

[14] J. Hertz, A. Krough, and R. G. Palmer. *Introduction to the Theory of Neural Computation.* Addison Wesley., 1993.

[15] M. Shiino and T. Fukai. Self-Consistent Signal-to-Noise Analysis and its Application to Analogue Neural Networks with Asymmetric Connections. *Journal of Physics A: Mathematical and General*, 25:L375–L381, 1992.

[16] S. Amari and K. Maginu. Statistical Neurodynamics of Associative Memory. *Neural Networks*, 1:63–73, 1988.

[17] M. Okada. A Hierarchy of Macrodynamical Equations for Associative Memory. *Neural Networks*, 8:833–838, 1995.

[18] D. J. Amit, H. Gutfreund, and H. Sompolinsky. Storing Infinite Numbers of Patterns in a Spin-glass Model of Neural Networks. *Physical Review Letters*, 55:1530–1533, 1985.

APPENDIX A: DETAIL SCSNA DESCRIPTION

In Sec.III, we introduced the overlaps,

$$m_\mu = \frac{1}{cN} \sum_{i=1}^{cN} \xi_i^\mu x_i \tag{A1}$$

$$\tilde{m}_\mu = \frac{1}{\tilde{c}N} \sum_{j=1}^{\tilde{c}N} \tilde{\xi}_j^\mu \tilde{x}_j \tag{A2}$$

for each layer state. In the equilibrium state, we assumed that the first pattern pair, ($\boldsymbol{\xi}^1$, $\tilde{\boldsymbol{\xi}}^1$), is retrieved, so that overlaps for other pattern pairs are small, *i.e.* $m_\mu, \tilde{m}_\mu \sim O(\frac{1}{\sqrt{N}})$ where $\mu \geq 2$. Thus we denotes the m_μ

$$m_\mu = \frac{1}{cN}\sum_{i=1}^{cN} \xi_i^\mu F(\tilde{c}\sum_{\mu=1}^{\alpha N} \xi_i^\mu \tilde{m}_\mu) \tag{A3}$$

$$\sim \frac{1}{cN}\sum_{i=1}^{cN} \xi_i^\mu F(\tilde{c}\sum_{\nu\neq\mu}^{\alpha N} \xi_i^\nu \tilde{m}_\nu) + \frac{1}{cN}\sum_{i=1}^{cN} \tilde{c}\tilde{m}_\mu F'(\tilde{c}\sum_{\nu\neq\mu}^{\alpha N} \xi_i^\nu \tilde{m}_\nu) \tag{A4}$$

$$= M_\mu + \tilde{c}\tilde{m}_\mu U, \tag{A5}$$

where

$$U = \frac{1}{cN}\sum_{i=1}^{cN} F'(\tilde{c}\sum_{\nu\neq\mu}^{\alpha N} \xi_i^\nu \tilde{m}_\nu) \tag{A6}$$

$$M_\mu = \frac{1}{cN}\sum_{i=1}^{cN} \xi_i^\mu F(\tilde{c}\sum_{\nu\neq\mu}^{\alpha N} \xi_i^\nu \tilde{m}_\nu) = \frac{1}{cN}\sum_{i=1}^{cN} \xi_i^\mu x_i^{(\mu)}. \tag{A7}$$

We denote the $x_i^{(\mu)}$ as the value drawn the effect of μth pattern pair from x_i, *i.e.* $x_i - x_i^{(\mu)} \sim O(\frac{1}{\sqrt{N}})$. $\tilde{m}_\mu$ is also denoted

$$\tilde{m}_\mu \sim \tilde{M}_\mu + cm_\mu\tilde{U}, \tag{A8}$$

where

$$\tilde{U} = \frac{1}{\tilde{c}N}\sum_{i=1}^{\tilde{c}N} F'(c\sum_{\nu\neq\mu}^{\alpha N} \tilde{\xi}_j^\nu m_\nu) \tag{A9}$$

$$\tilde{M}_\mu = \frac{1}{\tilde{c}N}\sum_{j=1}^{cN} \tilde{\xi}_j^\mu \tilde{x}_j^{(\mu)}. \tag{A10}$$

Solving eqs.(A5) and (A8) for m^μ and $\tilde{m}^\mu$, we obtain

$$m_\mu = \frac{1}{1-c\tilde{c}U\tilde{U}}(M_\mu + \tilde{c}U\tilde{M}_\mu), \tag{A11}$$

$$\tilde{m}_\mu = \frac{1}{1-c\tilde{c}U\tilde{U}}(\tilde{M}_\mu + c\tilde{U}M_\mu). \tag{A12}$$

Since the noise terms in eqs.(12) and (13) can be described as $z_i = \tilde{c}\sum_{\nu\geq 2}^{\alpha N} \xi_i^\nu \tilde{m}_\nu$, and $\tilde{z}_j = c\sum_{\nu\geq 2}^{\alpha N} \tilde{\xi}_i^\nu m_\nu$, we substituted eqs.(A11) and (A12) to these noises and obtained

$$z_i = \frac{\alpha\tilde{c}\tilde{U}}{1-c\tilde{c}U\tilde{U}} x_i^{(\mu)} + Z_i, \tag{A13}$$

$$\tilde{z}_j = \frac{\alpha c U}{1-c\tilde{c}U\tilde{U}} \tilde{x}_j^{(\mu)} + \tilde{Z}_j, \tag{A14}$$

where

$$Z_i = \frac{1}{N(1-c\tilde{c}U\tilde{U})}\sum_{\nu\geq 2}^{\alpha N}\left(\tilde{c}\tilde{U}\sum_{k\neq i}\xi_i^\nu\xi_k^\nu x_k^{(\mu)} + \sum_{j=1}^{\tilde{c}N}\xi_i^\nu\tilde{\xi}_j^\nu\tilde{x}_j^{(\mu)}\right) \tag{A15}$$

$$\tilde{Z}_j = \frac{1}{N(1-c\tilde{c}U\tilde{U})}\sum_{\nu\geq 2}^{\alpha N}\left(cU\sum_{l\neq j}\tilde{\xi}_j^\nu\tilde{\xi}_l^\nu\tilde{x}_l^{(\mu)} + \sum_{i=1}^{cN}\tilde{\xi}_j^\nu x i_i^\nu x_i^{(\mu)}\right) \tag{A16}$$

We assumed Z_i and $\tilde{Z}_j$ as independent identical Gaussian noise, described as $Z_i \sim N(0, \alpha r)$ and $\tilde{Z}_j \sim N(0, \alpha\tilde{r})$, and evaluated expectations $E[(Z_i)^2]$ and $E[(\tilde{Z}_j)^2]$. We then obtained

$$E[(Z_i)^2] = \alpha r = \frac{\alpha\tilde{c}}{(1-c\tilde{c}U\tilde{U})^2}(c\tilde{c}\tilde{U}^2 q + \tilde{q}), \tag{A17}$$

$$E[(\tilde{Z}_j)^2] = \alpha\tilde{r} = \frac{\alpha c}{(1-c\tilde{c}U\tilde{U})^2}(c\tilde{c}U^2\tilde{q} + q), \tag{A18}$$

where

$$q = \frac{1}{cN}\sum_{i=1}^{cN}(x_i^{(\mu)})^2 \tag{A19}$$

$$\tilde{q} = \frac{1}{\tilde{c}N}\sum_{j=1}^{\tilde{c}N}(\tilde{x}_j^{(\mu)})^2 \tag{A20}$$

From the self-averaging property, we obtained the SCSNA order parameter equation in sec.III.

APPENDIX B: CORRELATION WITH N-STEP PREVIOUS STATE

To evaluate the effect of the previous crosstalk noises, we must derive the correlation of a unit between the current state and the n-step before state. These are described by $\tilde{q}_{2t+1,2(t-n)+1}$ and $q_{2t,2(t-n)}$, respectively:

$$\begin{aligned}
&\tilde{q}_{2t+1,2(t-n)+1} = \\
&\frac{\int D\boldsymbol{z}\exp(-\boldsymbol{z}^{\mathrm{T}}\tilde{\Sigma}^{-1}\boldsymbol{z})\langle\tilde{Y}_{2t+1}(z_1)\tilde{Y}_{2(t-n)+1}(z_2)\rangle}{2\pi\left|\tilde{\Sigma}\right|} \\
&q_{2t,2(t-n)} = \\
&\frac{\int D\tilde{\boldsymbol{z}}\exp(-\tilde{\boldsymbol{z}}^{\mathrm{T}}\Sigma^{-1}\tilde{\boldsymbol{z}})\langle Y_{2t}(\tilde{z}_1)Y_{2(t-n)}(\tilde{z}_2)\rangle}{2\pi\left|\Sigma\right|}
\end{aligned} \tag{B1}$$

The matrices Σ, $\tilde{\Sigma}$ and vectors $\boldsymbol{z}, \tilde{\boldsymbol{z}}$ are described as follows:

$$\begin{aligned}
\tilde{\Sigma} &= \begin{pmatrix} \tilde{r}_{2t} & \tilde{r}_{2t,2(t-n)} \\ \tilde{r}_{2t,2(t-n)} & \tilde{r}_{2(t-n)} \end{pmatrix} \\
\boldsymbol{z} &= \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \\
\Sigma &= \begin{pmatrix} r_{2t-1} & r_{2t-1,2(t-n)-1} \\ r_{2t,2(t-n)} & r_{2(t-n)-1} \end{pmatrix} \\
\tilde{\boldsymbol{z}} &= \begin{pmatrix} \tilde{z}_1 \\ \tilde{z}_2 \end{pmatrix}
\end{aligned} \tag{B2}$$

The diagonal components of each matrix correspond to the variances of the current state noises. The non-diagonal components express the noise correlation between the current state and the η-step previous state. For $\eta \geq 2$, the correlations are described as:

$$\begin{aligned}
r_{2t+1,2(t-\eta)+1} =& c\tilde{c}\tilde{q}_{2t+1,2(t-\eta+1)+1} \\
&+ c\tilde{c}\tilde{U}_{2t+1}U_{2t}r_{2t-1,2(t-\eta+1)+1}, \\
\tilde{r}_{2t,2(t-\eta)} =& cq_{2t,2(t-\eta)} \\
&+ c\tilde{c}U_{2t}\tilde{U}_{2t-1}\tilde{r}_{2(t-1),2(t-\eta)},
\end{aligned} \tag{B3}$$

and for $\eta \geq 3$, they are

$$
\begin{aligned}
r_{2t+1,2(t-\eta)+1} &= \tilde{c}\tilde{q}_{2t+1,2(t-\lambda)+1} \\
&+ c\tilde{c}^2 \tilde{U}_{2t+1}\tilde{U}_{2(t-\lambda)+1} q_{2t,2(t-\lambda)} \\
&+ \tilde{c}\sum_{\lambda=1}^{n} (c\tilde{c})^{\lambda} \tilde{q}_{2(t-\lambda)+1,2(t-\eta)+1} \prod_{\tau=t-\lambda+1}^{t} \tilde{U}_{2\tau+1} U_{2\tau} \\
&+ \tilde{c}\sum_{\lambda=1}^{n-\eta} (c\tilde{c})^{\lambda} \tilde{q}_{2t+1,2(t-\eta-\lambda)+1} \prod_{\tau=t-\lambda+1}^{t} \tilde{U}_{2(\tau-\eta)+1} U_{2(\tau-\eta)} \\
&+ c\tilde{c}^2 \tilde{U}_{2t+1}\tilde{U}_{2(t-\eta)+1} \\
&\qquad \sum_{\lambda=1}^{n-1} (c\tilde{c})^{\lambda} q_{2(t-\lambda),2(t-\eta)} \prod_{\tau=t-\lambda+1}^{t} U_{2\tau}\tilde{U}_{2\tau-1} \\
&+ c\tilde{c}^2 \tilde{U}_{2t+1}\tilde{U}_{2(t-\eta)+1} \sum_{\lambda=1}^{n-1-\eta} (c\tilde{c})^{\lambda} q_{2t,2(t-\eta-\lambda)} \\
&\qquad \prod_{\tau=t-\lambda+1}^{t} U_{2(\tau-\eta)}\tilde{U}_{2(\tau-\eta)-1} \\
&+ (c\tilde{c})^2 \tilde{U}_{2t+1} U_{2t} \tilde{U}_{2(t-\eta)+1} U_{2(t-\eta)} r_{2t-1,2(t-\lambda)-1}
\end{aligned}
$$

$$
\begin{aligned}
\tilde{r}_{2t,2(t-\eta)} &= c q_{2t,2(t-\eta)} \\
&+ \tilde{c}c^2 U_{2t} U_{2(t-\eta)} \tilde{q}_{2t-1,2(t-\eta)-1} \\
&+ c\sum_{\lambda=1}^{n} (c\tilde{c})^{\lambda} q_{2(t-\lambda),2(t-\eta)} \prod_{\tau=t-\lambda+1}^{t} U_{2\tau}\tilde{U}_{2\tau-1} \\
&+ c\sum_{\lambda=1}^{n-\eta} (c\tilde{c})^{\lambda} q_{2t,2(t-\eta-\lambda)} \prod_{\tau=t-\lambda+1}^{t} U_{2(\tau-\eta)}\tilde{U}_{2(\tau-\eta)-1} \\
&+ \tilde{c}c^2 U_{2t} U_{2(t-\eta)} \\
&\qquad \sum_{\lambda=1}^{n-1} (c\tilde{c})^{\lambda} \tilde{q}_{2(t-\lambda)-1,2(t-\eta)-1} \prod_{\tau=t-\lambda+1}^{t} \tilde{U}_{2\tau-1} U_{2\tau-2} \\
&+ \tilde{c}c^2 U_{2t} U_{2(t-\eta)} \sum_{\lambda=1}^{n-1-\eta} (c\tilde{c})^{\lambda} \tilde{q}_{2t-1,2(t-\eta-\lambda)-1} \\
&\qquad \prod_{\tau=t-\lambda+1}^{t} \tilde{U}_{2(\tau-\eta)-1} U_{2(\tau-\eta)-2} \\
&+ (c\tilde{c})^2 U_{2t} \tilde{U}_{2t-1} U_{2(t-\eta)} \tilde{U}_{2(t-\eta)-1} \tilde{r}_{2(t-1),2(t-\lambda-1)}
\end{aligned}
\tag{B4}
$$

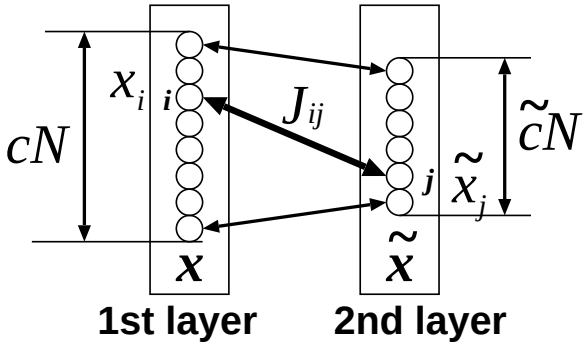


FIG. 1: Network structure of BAM

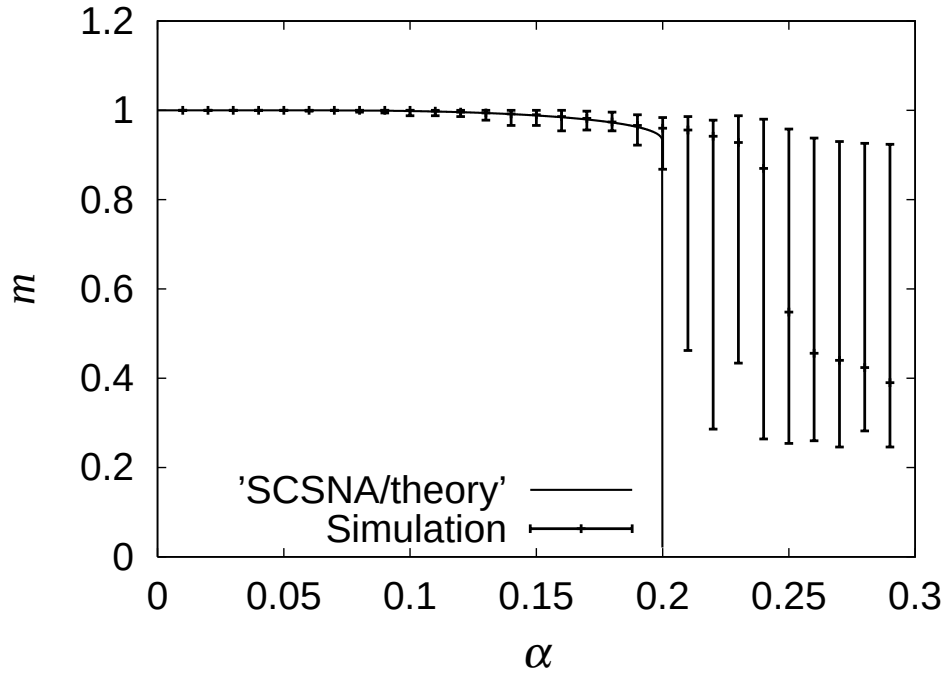
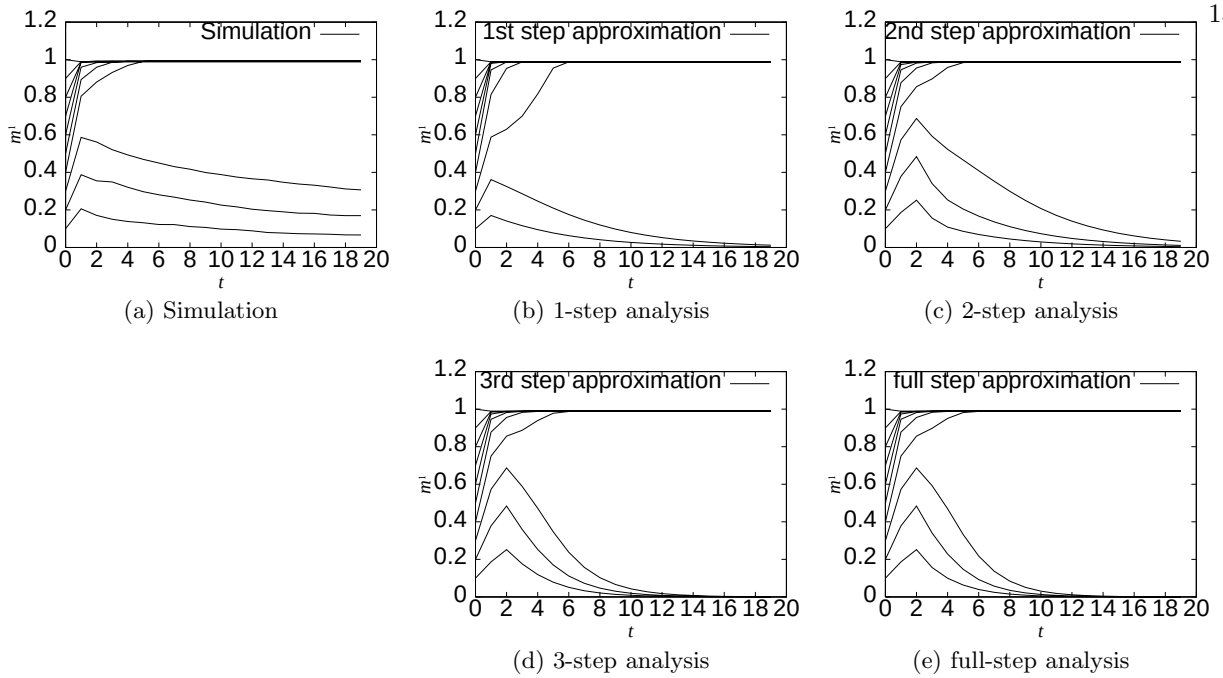


FIG. 2: Comparing SCSNA results with those of the simulation. The horizontal axis means the loading rate α, and the vertical axis means the overlap. The results of computer simulation are shown as error-bars, which indicates median with minimum and maximum values.



(a) Simulation (b) 1-step analysis (c) 2-step analysis

(d) 3-step analysis (e) full-step analysis

FIG. 3: Retrival process of a computer simulation and the statistical neurodynamics. The horizontal axis means time index t, and the vertical axis means the overlap m. (a) shows a result of computer simulation. From (b) to (e) shows the results of statistical neurodynamics.

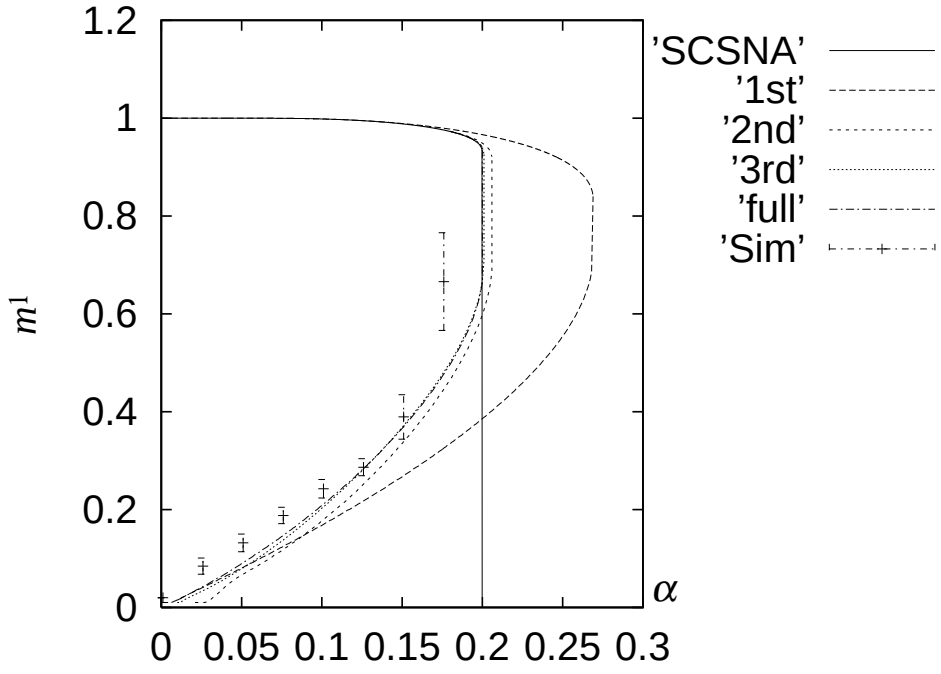


FIG. 4: Capacity comparing the statistical neurodynamics with SCSNA. The horizontal axis means the loading rate α, and the vertical axis means the overlap m. The dashed curves shows the results of the statistical neurodynamics. The results of computer simulations are shown with error-bar which indicates mean with standard deviations.